

Pengenalan Suara Burung Menggunakan *Mel Frequency Cepstrum Coefficient* dan Jaringan Syaraf Tiruan pada Sistem Pengusir Hama Burung

Fajar Budiman, Muhammad Agung Nursyeha, Muhammad Rivai, dan Suwito

Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember (ITS)

*Corresponding author, e-mail : fajarbudiman@ee.its.ac.id

Abstrak—Indonesia merupakan salah satu negara agraris yang memproduksi hasil pertanian. Namun, Indonesia masih mengimpor beras dari negara lain dikarenakan penurunan hasil panen. Salah satu faktor menurunnya produksi beras Indonesia adalah akibat serangan hama burung. Ekosistem sawah mengandung berbagai macam spesies burung, baik hama maupun non-hama. Burung non-hama menolong petani melawan hama serangga. Petani menggunakan metode tradisional untuk mengusir hama burung. Pada penelitian ini telah dirancang perangkat lunak untuk mengenali jenis burung berdasarkan kicauannya. *Voice Activity Detection* (VAD) digunakan untuk mendeteksi adanya kicau burung. Metode ekstraksi ciri suara dari kicau burung menggunakan *Mel Frequency Cepstrum Coefficient* (MFCC) dan *Fast Fourier Transform* (FFT). Jaringan Syaraf Tiruan digunakan untuk mengenali pola hasil ekstraksi. Selanjutnya, *audiosonic bird repeller* digunakan sebagai metode pengusiran hama burung. Hasil identifikasi *offline* dengan menggunakan MFCC didapatkan tingkat keberhasilan mencapai 90% untuk variasi kicauan dan jenis burung, sedangkan dengan FFT mencapai 68%. Hasil identifikasi *online* untuk spesimen burung bondol didapatkan tingkat keberhasilan 70% dengan menggunakan MFCC, dan 30% dengan FFT. Selain itu, suara tembakan merupakan suara yang paling baik digunakan untuk mengusir hama burung. Tingkat keberhasilan pengenalan suara burung menggunakan ekstraksi ciri MFCC lebih tinggi jika dibandingkan dengan ekstraksi ciri dengan menggunakan FFT.

Kata Kunci : Hama Burung, *Voice Activity Detection*, *Fast Fourier Transform*, dan *Mel Frequency Cepstrum Coefficient*.

Abstract—Indonesia is one of the agricultural country that produces crops. Nevertheless, Indonesia still imports rice from other countries because of crop decreasement. One of its factors is caused by bird pests. In the ricefields ecosystem, the variety of bird species can be classified as a pest and as a non-pest. Non-pest birds usually help farmer against insects. Farmers are using traditional methods to repel bird pests. In this research, a software to recognize species of birds has been designed. The system is based on birdchirp types. A mono-microphone is used to capture the sound. *Voice Activity Detection* (VAD) method is used for birdchirp detection. *Mel Frequency Cepstrum Coefficient* (MFCC) and *Fast Fourier Transform* (FFT) are used to extract birdchirp feature. Artificial Neural Network is utilized to recognize the pattern of birdchirp feature. Furthermore, *audiosonic bird repeller* is used to repel bird pests. In the offline mode testing, success level using MFCC feature extraction is up to 90% for birdchirp variation, while up to 68% using FFT. In the online mode, the average success level using MFCC feature extraction is 70% for finch birds, while 30% using FFT. In addition, a gunshot is a best sound to repel bird pests. The success rate of bird voice recognitions using MFCC feature extraction is higher than that using FFT.

Keywords : Bird pests, *Voice Activity Detection*, *Fast Fourier Transform*, *Mel Frequency Cepstrum Coefficient*.

Copyright © 2016 JNTE. All rights reserved

1. PENDAHULUAN

Indonesia merupakan salah satu negara agraris yang memproduksi hasil pertanian untuk kebutuhan pangan. Kebutuhan pangan masyarakat Indonesia dimungkinkan dapat dipenuhi dari hasil produksi secara mandiri.

Penduduk Indonesia mayoritas mengkonsumsi beras sebagai bahan pangan pokok. Produksi beras dalam negeri belum mampu mencukupi kebutuhan pangan dalam negeri, sehingga pemerintah Indonesia harus mengimpor beras dari negara lain [1]. Ketidakmampuan produksi beras dalam negeri untuk mencukupi kebutuhan

pangan disebabkan oleh penurunan produksi beras.

Salah satu faktor yang menyebabkan penurunan produksi beras di Indonesia adalah serangan hama burung. Beberapa jenis burung yang berada pada ekosistem sawah mengkonsumsi biji padi [2]. Pestisida untuk mengurangi populasi hama burung pada ekosistem sawah belum ditemukan, sehingga petani umumnya masih menggunakan cara konvensional untuk metode pengusiran hama burung. Salah satu metode konvensional pengusiran hama burung adalah dengan orang-orangan sawah. Operator atau petani akan menggerakkan orang-orangan sawah agar menimbulkan suara yang akan mengganggu hama burung. Kelemahan dari metode pengusiran hama burung ini adalah tidak mampu berjalan selama 24 jam, sehingga metode pengusiran ini kurang efisien. Oleh karena itu diperlukan teknologi yang dapat mengusir hama burung secara efisien.

Suara kicau burung dapat digunakan sebagai salah satu parameter pendeteksian keberadaan burung. Masing-masing spesies burung mempunyai kicau dan bunyi yang berbeda [3], sehingga jenis burung dapat diidentifikasi berdasarkan kicauannya.

Pada penelitian ini dirancang perangkat lunak untuk sistem pengusir hama burung yang mendeteksi jenis hama burung berdasarkan kicau burung. Diharapkan dengan adanya perangkat lunak ini, metode pengusiran hama burung pada ekosistem sawah dapat berjalan dengan lebih efisien.

2. TINJAUAN PUSTAKA

2.1. Burung pada Ekosistem Sawah

Sawah merupakan sebuah ekosistem buatan manusia, sehingga didalamnya terdapat komponen pembentuk ekosistem. Masing-masing komponen saling berhubungan dan memberikan timbal balik. Komponen pembentuk ekosistem sawah terdiri atas komponen biotik dan abiotik [2][4].

Komponen biotik terdiri atas makhluk hidup yang mendiami area persawahan. Sawah merupakan habitat beberapa makhluk hidup dikarenakan ketersediaan makanan. Salah satu contoh makhluk hidup yang menempati sawah adalah burung. Terdapat beberapa jenis burung yang menempati ekosistem persawahan. Adapun

burung yang menempati daerah persawahan antara lain burung cekakak jawa, burung bondol jawa, burung gereja, burung perkutut dan burung gelatik [2][5].

2.2. Voice Activity Detection

Voice Activity Detection (VAD) merupakan proses yang penting untuk pemrosesan sinyal suara. VAD digunakan pada awal pemrosesan sinyal suara. VAD merupakan metode untuk membedakan antara tutur dengan keadaan diam [6]. Metode VAD menganalisa energi dan frekuensi dominan dari sejumlah data suara [7][8]. Analisa VAD mengikuti persamaan (1) dan (2).

$$\phi_{power} = \frac{1}{N} \sum_{n=0}^{N-1} x[n]^2, 0 \leq n \leq N-1 \quad (1)$$

$$\phi_{frekuensi} = \max(FFT(x[n])) \quad (2)$$

dimana ϕ_{power} adalah energi sinyal suara pada sejumlah N sampel sinyal suara dan $\phi_{frekuensi}$ merupakan frekuensi dominan dari sejumlah N sampel sinyal suara yang telah mengalami transformasi *fourier*.

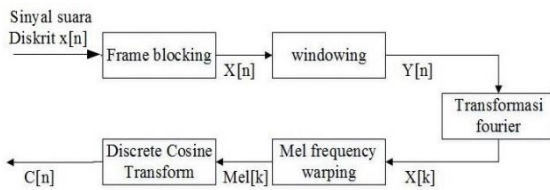
2.3. Mel Frequency Cepstrum Coefficient

Mel Frequency Cepstrum Coefficient (MFCC) merupakan metode ekstraksi ciri suara yang meniru cara kerja telinga manusia [9][10]. Bentuk sinyal suara konstan dalam waktu yang singkat antara 20ms hingga 30ms. Kemudian bentuk sinyal berubah. Perubahan bentuk sinyal terhadap waktu merupakan perbedaan bunyi yang dihasilkan dari organ penghasil vokal.

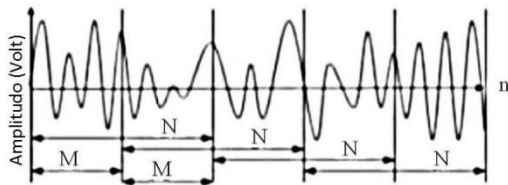
Telinga manusia sebagai organ penangkap suara bertindak sebagai filter. Filter yang bekerja merupakan filter *bandpass*. Lebar pita frekuensi telinga manusia bekerja secara linier pada frekuensi dibawah 1000Hz dan melebar secara logaritmik pada frekuensi lebih dari 1000Hz [10].

Proses MFCC terdiri atas *frame blocking*, *windowing*, *fast fourier transform*, *mel frequency warping* dan *discrete cosine transform*. Secara garis besar diagram blok dari proses MFCC dapat divisualisasikan seperti terlihat pada Gambar 1.

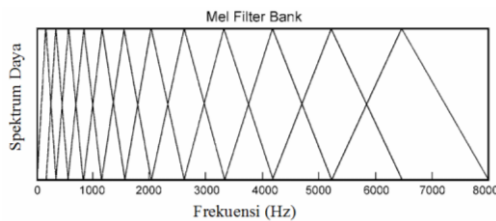
2.3.1. Proses *Frame Blocking*



Gambar 1. Diagram Blok MFCC



Gambar 2. Proses *Frame Blocking*



Gambar 3. Bentuk *Mel Filter Bank*

Pada proses *frame blocking*, analisa suara yang dilakukan berupa *short time analysis*. Lebar waktu analisa berkisar antara 20ms hingga 30ms. Sinyal suara $x[n]$ memiliki durasi yang panjang dibagi menjadi sejumlah *frame*. Satu buah *frame* memiliki N sampel data suara. Antara satu *frame* dengan *frame* yang lain saling *overlapping* sejumlah M sampel data suara dimana nilai M tidak lebih dari nilai N [11]. Proses *frame blocking* dapat divisualisasikan seperti pada Gambar 2.

2.3.2. Proses *Windowing*

Windowing merupakan pendekatan filter digital *Finite Impulse Response* (FIR) [12]. Proses ini menghilangkan sinyal *aliasing* akibat diskontinuitas sinyal. Diskontinuitas terjadi diakibatkan oleh proses *frame blocking* [10][11]. Fungsi *window* yang baik menghasilkan bentuk sinyal yang runcing pada *main lobe* dan teredam pada *side lobe*, yang dituangkan pada persamaan (3).

$$y(n) = x(n)w(n), 0 \leq n \leq N-1 \quad (3)$$

dimana $y(n)$ adalah sinyal hasil dari konvolusi antara sinyal masukan dan fungsi *window*, $x(n)$ merupakan sinyal yang akan dikonvolusikan dengan fungsi *window*.

2.3.3. Proses *Fast Fourier Transform*

Suatu fungsi dengan periode terbatas dapat dinyatakan dalam deret *fourier*. Transformasi *fourier* digunakan untuk mengubah deret sinyal domain waktu terbatas menjadi spektrum frekuensi [10][11]. *Frame* yang telah mengalami proses *windowing* diubah menjadi spektrum frekuensi. Sedangkan, *Fast Fourier Transform* (FFT) merupakan algoritma cepat dari *Discrete Fourier Transform* (DFT). FFT mengurangi perkalian berulang yang terdapat pada DFT, yang tertuang pada persamaan (4).

$$x[k] = \sum_{n=0}^{N-1} y[n] * e^{-2\pi jkn/N}, n=0,1,...,N-1 \quad (4)$$

Dimana $x[k]$ merupakan pola frekuensi ke- k yang dihasilkan dari transformasi *fourier*, $y[n]$ adalah sinyal dari sebuah *frame*.

2.3.4. Proses *Mel Frequency Warping*

Persepsi telinga manusia terhadap frekuensi suara tidak mengikuti skala linier [10][11]. Skala frekuensi sebenarnya menggunakan satuan Hz. Skala yang bekerja pada telinga manusia disebut skala mel frekuensi, yang dirumuskan seperti persamaan (5).

$$mel(k) = 2595 * \log_{10} \left(1 + \frac{f}{700} \right) \quad (5)$$

Dimana $mel(k)$ merupakan skala mel ke- k dan f merupakan frekuensi dalam Hz. Salah satu pendekatan spektrum frekuensi dalam skala mel dengan fungsi kerja telinga manusia sebagai filter adalah dengan *filter bank*. *Filter bank* merupakan kumpulan filter yang dapat divisualisasikan dalam Gambar 3.

2.3.5. Proses *Discrete Cosine Transform* (DCT)

Manusia mendengarkan informasi suara berdasarkan sinyal pada domain waktu. *Discrete Cosine Transform* (DCT) digunakan untuk mengubah spektrum frekuensi kedalam *cepstrum*. *Cepstrum* merupakan kebalikan dari

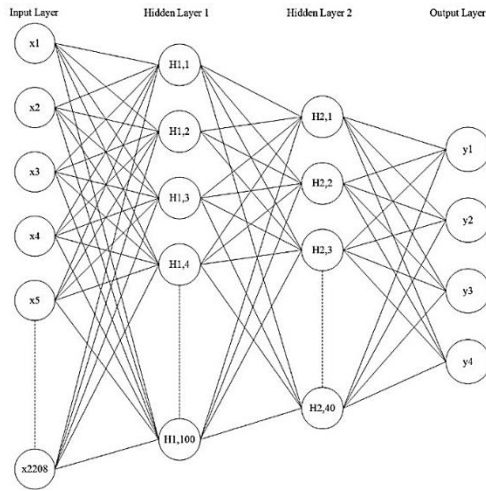
spektrum [10][11]. Ciri fonetis dari sinyal suara tidak hilang setelah dikembalikan kedalam bentuk *cepstrum*. Akhirnya, koefisien MFCC bisa dirumuskan seperti pada persamaan (6).

$$C_n = \sum_{k=1}^K (\log S_k) \cos \left[n \left(k - \frac{1}{2} \right) \frac{\pi}{K} \right] \quad (6)$$

dimana C_n merupakan koefisien MFCC, S_k merupakan *power spectrum* mel frekuensi dan K jumlah *filter bank*.

2.4. Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan (JST) merupakan kecerdasan buatan yang meniru sistem syaraf manusia. JST merupakan metode pengenalan pola [13]. Pembelajaran JST berupa perubahan nilai bobot yang terdapat dalam arsitektur JST. Bobot saling menghubungkan antara masing-masing noda *input* dan *neuron* yang terdapat pada arsitektur JST. Arsitektur JST yang digunakan pada sistem ini seperti Gambar 4.



Gambar 4. Arsitektur Jaringan Syaraf Tiruan

Secara umum dalam arsitektur Jaringan Syaraf Tiruan terdapat noda *input*, *neuron*, *hidden layer*, *output layer* dan bobot. Algoritma pembelajaran dalam satu iterasi pada jaringan syaraf tiruan terdapat 3 prosedur. Prosedur tersebut antara lain.

2.4.1. Feedforward

Dalam prosedur *feedforward*, *input* dikalikan dengan bobot yang terhubung dengan *neuron* lapisan berikutnya. Kemudian hasil dari proses penjumlahan dilewatkan fungsi aktivasi.

Prosedur *feedforward* digunakan untuk melihat respon *output* terhadap pola masukan.

$$\varphi_i = \sum_{j=0}^J x[j] * w[j, i] \quad (7)$$

dimana φ_i adalah noda atau neuron ke- i pada suatu *layer*, $x[j]$ merupakan input atau *neuron* pada *layer*, $w[j, i]$ adalah bobot dari input ke j menuju *neuron* ke- i .

2.4.2. Backpropagation

Dalam prosedur *backpropagation* dapat diketahui galat antara pola keluaran dengan pola data keluaran yang diinginkan [14]. Untuk mengetahui besar galat pada *output layer* digunakan *Mean Square Error* (MSE) yang dihitung sesuai persamaan (8). Sedangkan untuk mengetahui galat pada masing-masing *neuron* digunakan persamaan (9).

$$MSE = \frac{1}{N} \sum_{n=1}^N (d_i - \chi_i)^2 \quad (8)$$

$$\delta_i = \left(\sum_{j=1}^J \delta_j w_{ij} \right) g(\varphi_i) \quad (9)$$

Dimana d_i merupakan target pembelajaran, δ_j merupakan nilai keluaran aktual, δ_i merupakan galat pada suatu neuron dan $g(\varphi)$ turunan fungsi aktivasi pada neuron ke- i .

2.4.3. Update weight

Prosedur yang digunakan untuk memperbarui nilai bobot. Pembaruan bobot digunakan sebagai respon pembelajaran JST terhadap pola data masukan.

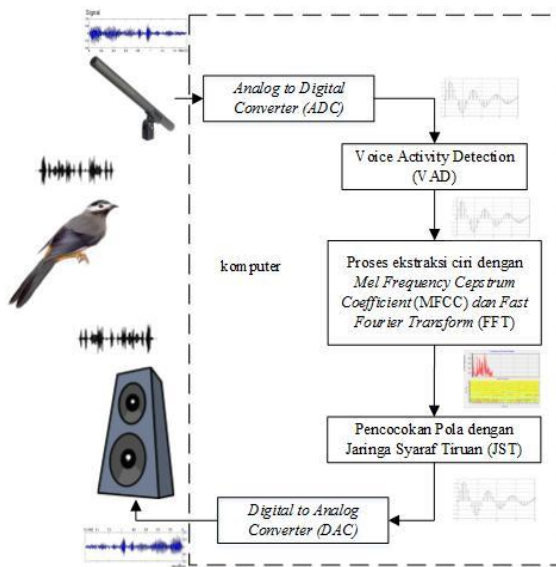
$$w_{baru}[j, i] = w[j, i] + \mu \delta_j x_i \quad (10)$$

Dimana w_{baru} merupakan bobot yang telah diperbarui, w merupakan bobot lama dan μ koefisien pembelajaran.

Setelah bobot diperbarui dalam suatu iterasi. Jumlah iterasi ditambah dan prosedur 1, 2 dan 3 diulang lagi hingga nilai galat MSE yang dihasilkan dari proses pembelajaran sesuai dengan galat MSE yang diinginkan.

2.5. Audiosonic Bird Repeller

Organ pendengaran makhluk hidup dipengaruhi oleh koklea. Pada burung, bentuk koklea lurus dan sedikit membengkok [15]. Hal ini mempengaruhi frekuensi suara yang mampu diterima oleh burung. Rata-rata burung merespon suara diantara rentang frekuensi 1kHz hingga 4kHz [15]. Rentang suara ini biasa disebut *audiosonic*.



Gambar 5. Diagram Blok Sistem Keseluruhan

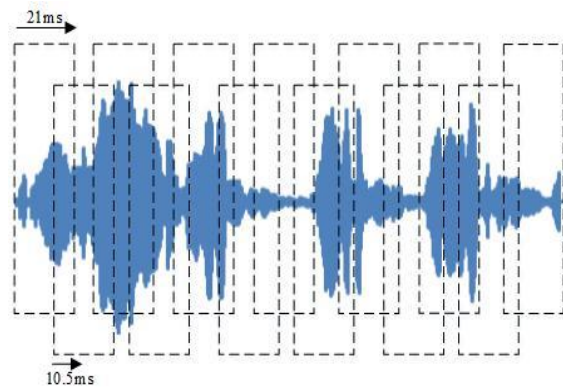
Salah satu metode pengusiran burung dengan menggunakan suara *audiosonic*. Suara yang direspon umumnya berupa suara yang merupakan informasi bagi burung. Suara tersebut dapat berupa suara dari predator burung [15].

3. PERANCANGAN SISTEM

Sistem yang dirancang memiliki fungsi untuk mengenali dan membedakan jenis burung berdasarkan bunyi kicauan. Diagram blok sistem dapat divisualisasikan pada Gambar 5. Cara kerja sistem yang dirancang secara keseluruhan adalah sebagai berikut :

1. Batas VAD ditentukan dengan durasi dua detik.
2. Batas VAD yang didapatkan berupa energi suara saat kondisi diam dan frekuensi dominan saat kondisi diam.
3. Durasi kicau yang terekam oleh VAD adalah satu detik.

4. Sistem pemrosesan akan mengekstraksi suara dengan metode MFCC dan FFT.
5. Hasil dari *cepstrum* MFCC atau spektrum frekuensi diidentifikasi menggunakan jaringan syaraf tiruan.
6. Apabila sistem mendeteksi jenis hama burung. Maka sistem merespon dengan memproduksi gelombang suara pengganggu.



Gambar 6. Proses Frame Blocking

Frekuensi pencuplikan sinyal suara yang digunakan adalah 48kHz. Dalam penelitian ini digunakan besar nilai N dan M masing-masing 1024 dan 512. Sehingga durasi dalam satu *frame* berdurasi 21ms dan bergeser 10.5ms yang dapat ditampilkan dalam Gambar 6. Daerah *overlap* antara dua *frame* adalah 50%.

Secara khusus dalam penelitian ini digunakan fungsi *window hamming*. Persamaan fungsi *window hamming* ditunjukkan pada persamaan (11). Persamaan *window hamming* menghasilkan spektrum frekuensi yang dominan pada *main lobe* dan teredam pada *side lobe*.

$$w(n) = 0.54 - 0.46 \left(\frac{2\pi n}{N-1} \right) \quad (11)$$

Jumlah *Mel filter bank* yang digunakan berjumlah 24 buah filter. Proses DCT menghasilkan 24 buah *cepstrum coefficient* dalam satu *frame*. Dalam durasi satu detik kicauan dihasilkan 92 *frame*. *Input* masukan jaringan syaraf tiruan merupakan perkalian antara jumlah koefisien *cepstrum* dan jumlah *frame*, sehingga didapatkan *input* jaringan syaraf tiruan sebanyak 2208 yang dapat divisualisasikan dalam Gambar 4.

4. HASIL PENGUJIAN

4.1. Voice Activity Detection

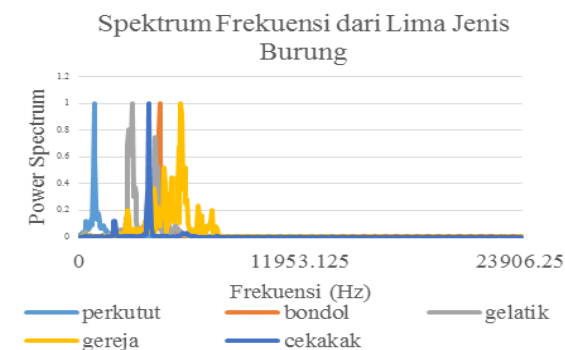
Dalam Tabel 1 ditunjukkan hasil dari pengujian *Voice Activity Detection* (VAD). Pengujian ini bertujuan untuk mengetahui batas yang dihasilkan oleh algoritma *Voice Activity Detection* (VAD) dalam beberapa kondisi akustik yang berbeda.

Tabel 1. Batas VAD Pada Beberapa Kondisi Akustik Berbeda

No	Kondisi	Frekuensi (Hz)		Energi (dB)	
		Atas	Bawah	Atas	Bawah
1	Senyap	656.25	229.21	36.5	33.97
2	Ruang Terbuka	703.12	246.56	41.2	34.26
3	Pinggir Jalan	562	237	57	51

Tabel 2. Kekuatan dan Frekuensi Dominan Suara Burung

No	Jenis burung	Frekuensi (Hz)	Kekuatan Suara (dB)		
			10cm	50cm	100cm
1	Gelatik	2813.9	82.5	60.4	51.4
2	Cekakak	3508.7	81.8	63.4	52.1
3	Bondol	3413.8	82.1	60.6	53
4	Gereja	3652.7	80.7	59.9	51.6
5	Perkutut	742	76.9	57.9	47.1



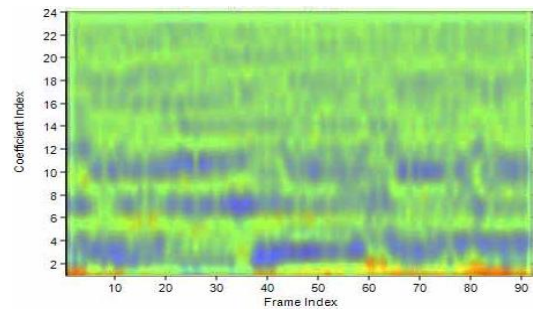
Gambar 7. Spektrum Frekuensi Burung yang Dideteksi

Kekuatan suara yang diterima oleh mikrofon dipengaruhi oleh jarak antara sumber kicauan dan mikrofon. Frekuensi dari kicau burung berbeda-beda. Untuk mengetahui kicau burung dapat dideteksi oleh VAD, digunakan nilai kekuatan kicauan dan frekuensi dominan dari kicau burung. Hasil dari pengujian didapatkan nilai pada Tabel 2.

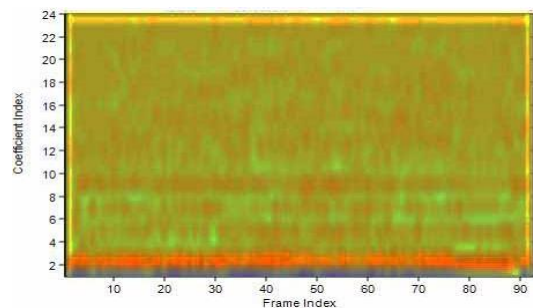
4.2. Pengujian Fast Forier Transform

Pengujian FFT bertujuan untuk melihat spektrum frekuensi dari kicau burung yang dapat ditunjukkan dalam Gambar 7. Masing-masing jenis burung yang diidentifikasi memiliki frekuensi kicau yang berbeda. Spektrum frekuensi yang diambil berupa *short time analysis* yang memiliki durasi 21 ms.v.

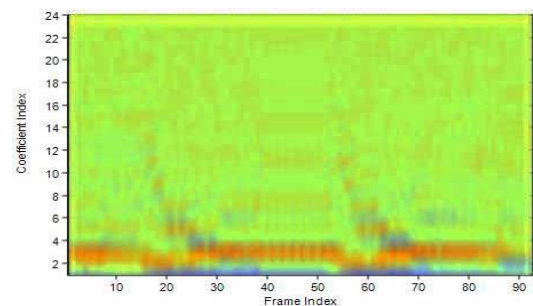
4.3. Pengujian Cepstrum MFCC



Gambar 8. Cepstrum MFCC burung Perkutut



Gambar 9. Cepstrum MFCC burung Gereja

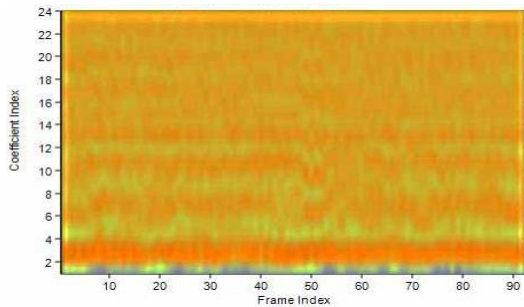


Gambar 10. Cepstrum MFCC burung Gelatik

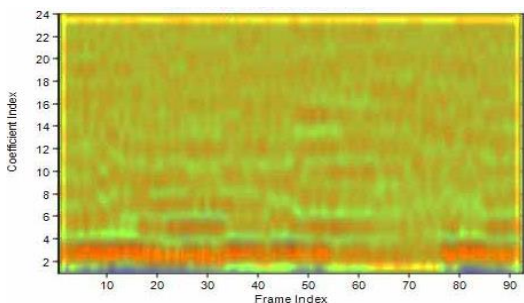
Pengujian ini bertujuan untuk melihat *cepstrum* dari kicau burung. Dengan menggunakan parameter *frame blocking* $N = 1024$ dan $M = 512$. Dengan frekuensi pencuplikan 48kHz, nilai N setara dengan 21ms dan M setara dengan 10.5ms. nilai N dan M

ditentukan agar mengurangi hilangnya karakteristik fonetis dari sinyal kicau akibat proses akuisisi data dengan VAD. Berdasarkan hasil pengujian yang ditunjukkan Gambar 8 hingga Gambar 12, *surface cepstrum* kicauan burung berbeda-beda. Perbedaan dapat dilihat dari kerapatan gradasi warna pada garis *coefficient index*.

4.4. Pengujian Offline dan Online



Gambar 11. *Cepstrum* MFCC burung Cekakak



Gambar 12. *Cepstrum* MFCC burung Bondol

Dari spektrum frekuensi dan *cepstrum* MFCC yang didapat dijadikan basis data. Kemudian dipelajari oleh Jaringan Syaraf Tiruan. Prosedur pengujian *offline* dilakukan dengan menggunakan berkas suara kicau burung. Berkas suara didapat dari *internet*. Variasi dalam pengujian *offline* berupa jarak antara pemutar audio dengan mikrofon, dan lokasi pengujian. Hasil pengujian *offline* disajikan dalam Tabel 3.

Pada pengujian *online* dilakukan agar pengenalan burung dapat diterapkan pada spesimen sebenarnya. Digunakan satu jenis burung dalam pengujian *online*. Jenis burung yang digunakan adalah burung bondol, dimana burung ini merupakan hama padi yang sangat banyak di sawah. Tingkat keberhasilan dari pengujian ini ditampilkan pada Tabel 4.

Suara pengusiran yang diujikan dalam penelitian ini adalah suara burung elang, suara kucing, suara tembakan dan suara kaleng. Prosedur pengujian dengan mengikat kaki spesimen dengan tali. Kemudian spesimen ditakuti dengan suara pengusir. Diamati tingkah laku burung. Dari hasil pengujian didapatkan keberhasilan pengusiran paling baik dengan menggunakan suara tembakan.

5. KESIMPULAN

Pengenalan suara burung dapat ditinjau berdasarkan pola *cepstrum* MFCC dan spektrum frekuensi. Dalam pengujian baik *online* maupun *offline* menunjukkan hasil pengenalan suara burung dengan spektrum frekuensi lebih rendah dibandingkan dengan menggunakan *cepstrum* MFCC.

Tabel 3. Pengujian *Offline* Sistem

No	Lokasi	Jarak (cm)	Keberhasilan (%)	
			MFCC	FFT
1	Outdoor (Taman)	10cm	84	60
		50cm	72	46
		100cm	52	38
		300cm	52	38
		500cm	34	42
2	Indoor	10cm	90	68
		50cm	78	48
		100cm	60	38
		300cm	54	34
		500cm	50	32
3	Outdoor (Pinggir Jalan)	10cm	44	44
		50cm	24	34
		100cm	0	0
		300cm	0	0
		500cm	0	0
4	Outdoor (Ramai, Pasar Burung)	10cm	50	52
		50cm	26	36
		100cm	14	28
		300cm	24	12
		500cm	30	24

Tabel 4. Pengujian *Online* Sistem

No	Jenis Burung	Keberhasilan (%)	
		MFCC	FFT
1	Bondol Jawa	60	30
		60	27
		80	30
		81	33
		60	33

Hal ini dikarenakan pengenalan dengan menggunakan spektrum frekuensi menggunakan *short time analysis* berdurasi 21ms. Dalam waktu yang singkat, spektrum frekuensi kicauan yang dihasilkan berbeda-beda. Keberhasilan pengenalan suara dipengaruhi oleh durasi kicau burung yang terekam oleh VAD. Tingkat keberhasilan pengenalan suara burung dengan menggunakan *cepstrum* MFCC mencapai 90% untuk variasi kicau burung dan jenis burung. Tingkat keberhasilan pengenalan suara burung dengan menggunakan spektrum frekuensi mencapai 68% untuk variasi kicau burung dan jenis burung. Untuk pengenalan suara burung bondol secara *online* didapatkan tingkat keberhasilan berkisar antara 60% hingga 80% dengan MFCC dan 27% hingga 33% dengan spektrum frekuensi. Suara tembakan merupakan suara yang baik untuk mengusir hama burung. Ekstraksi ciri suara kicau burung memberikan tingkat keberhasilan identifikasi yang lebih tinggi dibandingkan dengan ekstraksi ciri suara dengan FFT waktu pendek.

DAFTAR PUSTAKA

- [1] Nor Ain Firdaus, "Alat Pelacak Arah Suara pada Sistem Pengusir Hama Burung Menggunakan ARM STM32F4", Tugas Akhir Institut Teknologi Sepuluh Nopember, 2015.
- [2] Zainal Husain, Dharmono dan Kaspul, "Jenis dan Kerapatan Burung di Kawasan Agropolitan Kecamatan Mandastana Kabupaten Barito Kuala", Jurnal Wahana-Bio, 2010.
- [3] Alex L. McIlraith, H.C. Card, "Birdsong Recognition with DSP and Neural Network", IEEE Wescanex Proceeding. 1995.
- [4] D, Aryulina, et al, "Biologi SMA Untuk Kelas X", Esis, Jakarta, 2004.
- [5] Chrystanto, Siti Asyatun, Margareta R., "Keanekaragaman Avifauna di Cagar Alam Keling II/III Kabupaten Jepara Jawa Tengah", Indonesian Journal of Conservation Vol. 3 No. 1, 2014.
- [6] Elias Nemer, Rafik Goubiran, Samy Mahmoud, "Robust Voice Activity Detection Using Higher-Order Statistic in the LCP Residual Domain", IEEE Transaction on Speech and Audio Processing Vol. 9 No. 3, 2001.
- [7] M.H. Moattar, M.M. Homayounpaur, "A Simple but Efficient Real-Time Voice Activity Detection Algorithm", 17th European Signal Processing Conference, 2009.
- [8] Venkatesha R., Abhijeet Sangwan, H.S. Jamadagni, Chiranth M.C., Rahul Sah, "Comparison of Voice Activity Detection Algorithms for VoIP", Proceedings of the 7th International Symposium on Computers and Communication, 2002.
- [9] Fitrilina, Rahmadi Kurnia, dan Siska Aulia, "Pengenalan Ucapan Metoda MFCC-HMM untuk Perintah Gerak Robot Mobil Penjejak Identifikasi Warna", Jurnal Nasional Teknik Elektro Andalas, Vol. 2 No. 1, 2013.
- [10] Jorge Martinez, Hector Perez, Enrique Escamillia, Masahisa Mabo Suzuki, "Speaker Recognition Using Mel Frequency Cepstral Coefficient (MFCC) and Vector Quantization (VQ) Techniques", 22nd International Conference Electrical Communication and Computers (CONIELECOMP), 2012.
- [11] H.B. Kekre, V.A. Bharadi, A.R. Sawant, Onkar Kadam, Pushkar Lanke, Rohit Lodhiya, "Speaker Recognition using Vector Quantization by MFCC and KMCG Clustering Algorithm", International Conference on Communication, Information and Computing Technology (ICCICT), 2012.
- [12] Rohit Patel, Er. Mukesh Kumar, A.K. Jaiswal, Romini Saxena, "Design Technique of Bandpass FIR Filter using Various Window Function", IOSR Journal of Electronics and Communication Engineering, 2013.
- [13] Robert E. Uhrig, "Introduction to Artificial Neural Network", Proceedings of the 1995 IEEE IECON 21st Vol. 1, 1995.
- [14] M.N.H. Siddique, M.O. Tokhi, "Training Neural Networks: Back-propagation vs Genetic Algorithm", Proceedings IJCNN Vol.4, 2001.
- [15] Robert C. Beason, "What Can Birds Hear?", Proceeding 21st Vertebrata Pest Conference, 2004.

Biodata Penulis

Fajar Budiman, lahir di Sumenep, 7 Juli 1986. Penulis menamatkan pendidikan S1 di Teknik Elektro Institut Teknologi Sepuluh Nopember (ITS) pada tahun 2010 dan meneruskan S2 di Electronic and Computer Engineering Department di National Taiwan University of Science and Technology (NTUST) pada tahun 2011-2013. Pada saat ini, penulis menjadi dosen di jurusan Teknik Elektro, FTI ITS. Bidang penelitian yang digeluti berkisar pada Teknik Elektronika.

Muhammad Agung Nursyeha, lahir di Surabaya, 28 Maret 1992. Pada tahun 2011, penulis memulai pendidikan tinggi di jurusan Teknik Elektro ITS Surabaya. Selama kuliah penulis aktif sebagai asisten laboratorium Elektronika Dasar, serta aktif dalam Unit Kegiatan Mahasiswa ITS.

Muhammad Rivai, lahir pada tanggal 26 April 1969. Dari Maret 1994 hingga saat ini, penulis menjadi dosen Teknik Elektro ITS dan saat ini sekaligus menjabat sebagai Kepala Laboratorium Elektronika Industri. Bidang keahlian yang digeluti adalah bidang elektronika industri, sensor, teknologi sensor aroma, dan hidung elektronik.

Suwito, lahir di Sragen, 5 Januari 1981. Penulis menempuh pendidikan tingkat sarjana di Teknik Elektro ITS tahun 1999-2003. Kemudian, penulis sempat bekerja di Industri elektronik selama 2 tahun, dan kemudian menjadi dosen di ITS dari tahun 2005-sekarang. Penulis menyelesaikan pendidikan magister di ITB dari tahun 2007-2009 di bidang Industrial Automation, yang hingga sekarang bidang yang digeluti meliputi instrumentasi dan automasi industri.